

## ANALISA CIRI GENDER BERDASARKAN TIME DOMAIN AUDIO FEATURE

*Hari Purwadi<sup>1</sup>, Didi Susilo Budi Utomo<sup>2</sup>, Agusma Wajiansyah<sup>3</sup>, Tatik Yuliana<sup>4</sup>*

Program Studi Teknik Informatika Multimedia, Jurusan Teknologi Informasi, Politeknik Negeri Samarinda

e-mail: [1hari.purwa06@gmail.com](mailto:hari.purwa06@gmail.com), [2dsbudiutomo10@gmail.com](mailto:dsbudiutomo10@gmail.com), [3agusma.wajiansyah@gmail.com](mailto:agusma.wajiansyah@gmail.com),  
[4tatikyuliana13@gmail.com](mailto:tatikyuliana13@gmail.com)

### ABSTRAK

Sistem pengenalan pola merupakan komponen penting dalam proses peniruan kemampuan inderawi manusia terutama penglihatan dan pendengaran. dalam meniru indera pendengaran manusia, komputer harus mempunyai suatu mekanisme standar dan logis dalam mengenali pola yang ada pada suara yang sedang diproses. konsep sederhana untuk mengenali pola dari suara sehingga dapat diidentifikasi dengan baik oleh komputer [2]. Zero Crossing Rate merupakan Sempel berurutan pada sebuah signal digital memiliki perbedaan tanda, ukuran dari noise sebuah sinyal pada fitur domain. Sort Time Eney merupakan Eney menandakan kekerasan suara pada waktu yang pendek. Entropy merupakan Entropi adalah pemerosesan sinyal digital untuk mengkarakterisasi spectrum audio. Rerata wanita ZCR : 7366.4, STE: 0.0322, Entropy: 8.6222. Rerata pria ZSC: 6012.2, STE: 0.025, Entropy: 8.2924. Pada hasil rata-rata, nilai ciri suara ZCR tertinggi adalah 8615, dan ciri suara ZCR Terendah adalah 4266. nilai ciri suara STE tertinggi adalah 0.064, dan ciri suara STE Terendah adalah 0.013. nilai ciri suara Entropy tertinggi adalah 9.906, dan ciri suara Entropy Terendah adalah 8.135.

**Kata Kunci:** Sistem Pengenalan Pola, Suara, Zero Crossing Rate, Sort Time Energy, Entropy

### 1. PENDAHULUAN

Sistem pengenalan pola merupakan komponen penting dalam proses peniruan kemampuan inderawi manusia terutama penglihatan dan pendengaran. dalam meniru indera pendengaran manusia, komputer harus mempunyai suatu mekanisme standar dan logis dalam mengenali pola yang ada pada suara yang sedang diproses. konsep sederhana untuk mengenali pola dari suara sehingga dapat diidentifikasi dengan baik oleh komputer [3][2][4].

Pada penelitian Dilak (2017) "Klasifikasi Musik Menggunakan Polynomial Neural Network", Spectral Centroid (SC), Sort Time Eney (STE) dan Zero Crossing Rate (ZCR) dengan menggunakan Neural Network, menunjukkan bahwa pendekatan mampu melakukan klasifikasi terhadap jenis music dengan akurasi sebesar 90% [5][6][7].

Pada penelitian Dimasatria, virgono, dan Rumani M (2016) "Desain Implementasi Automatic Video Captioning dengan Speech Recognition Menggunakan Hidden Markov Model", hasil pengujian akurasi tertinggi Sebesar 75,50% pada jumlah State 6 dengan jumlah cluster 256 menggunakan data latik 90 untuk setiap suku ata [8][5]. Tujuan komputer dapat mengklasifikasikan gender dari hasil penentuan ciri Sort Time Eney (STE) dan Zero Crossing Rate (ZCR) dan Entropi.

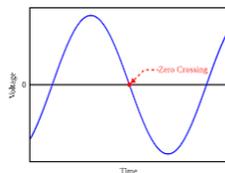
Dalam penelitian ini, ciri dari gender atau jenis kelamin akan diidhitung dalam domain waktu berdasarkan nilai ZCR, STE dan entropy untuk menentukan feture dari gender. Data ciri dari gender didapatkan akan berguna untuk penelitian lanjutan yang terkait dengan voice processing [9] [10]

### 2. METODE PENELITIAN

Pada penenitian ini dibagei menjadi beberapa bagian proses untuk mengetahui sebuah fitur gender dalam domain waktu diantaranya :

#### 2.1 Zero Crossing Rate

Zero-Crossing Rate: Zero-crossing rate (ZCR) adalah tingkat transisi sinyal dari positif ke nol ke negatif atau negatif ke nol ke positif [8] [11] [12].



Gambar 1. Zerro Crossing Rate

Sampel berurutan pada sebuah signal digital memiliki perbedaan tanda, ukuran dari noise sebuah sinyal pada fitur domain [8] [3] [13].

$$ZCR = \frac{1}{N} \sum_{n=1}^N |sign(x(n)) - sign(x(n-1))| \quad (1)$$

Keterangan:

ZCR = Zero crossing rate

Sign  $x(n)$  = Nilai dari  $x(n)$ , bernilai 1 jika  $x(n)$  positif, -1 jika  $x(n)$  negatif

N = Jumlah sampel

## 2.2 Sort Time Energy

Short Time Energy menandakan kekerasan suara pada waktu yang pendek, STE merupakan cara yang sederhana dan efektif untuk mengklasifikasikan parameter yang mengandung segmen bersuara atau tidak bersuara. STE juga digunakan untuk mendeteksi titik akhir ucapan [8] [3][14][15].

$$STE = \frac{1}{n} \sum_{n=1}^n x(n)^2 \quad (2)$$

Keterangan :

STE = Short time enetgy

N = Jumlah Sampel

X(n) = Nilai signal dari sampel

## 2.3 Entropy

Entropi adalah pemerosesan sinyal digital untuk mengkarakterisasi *spectrum audio*[16] [17]. Nilai entropy menunjukkan keteracakan distribusi derajat keabuan suat citra. Semakin acak distribusi derajat keabuanya, semakin tinggi nilai entropy yang dihasilkan

$$e = - \sum_{i=1}^n p(x_1) \log P(x_1) \quad (3)$$

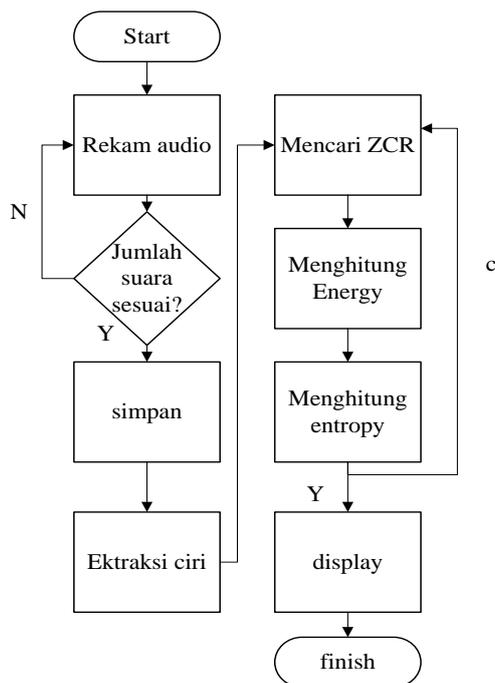
Keterangan :

e = Entropy

n = Jumlah Sampel

$P(x_i)$  = Probabilitas yang di dapat

Dari beberapa fitur yang akan dilakukan maka penelitian ini menggunakan metodologi sebagai berikut :



Gambar 2. Metodologi Penelitian

Pada gambar 2 merupakan proses-proses penelitian, proses awal dengan cara pengambilan data dari 10 responden, 5 wanita dan 5 pria. Dari hasil rekaman di simpan, data dari hasil rekaman dari 10 responden, dengan menghitung ZCR, STE, Entropy dengan menggunakan matlab, dari hasil perhitungan maka dapat hasil dari karakteristik dari masing-masing suara.

### 3. HASIL DAN PEMBAHASAN

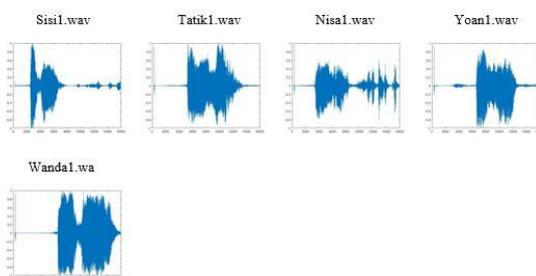
Pada bagian ini akan di jelaskan tahapan – tahapan penelitian sebagai berikut :

#### 3.1 Akusisi Data

Akusisi suara menggunakan *microfon* dengan menggunakan aplikasi Matlab [9] [18], data yang di simpan sudah format *wav*. suara yang direkam adalah suara dari lima orang wanita dan lima orang pria dengan masing-masing orang sebanyak 3 suara rekaman, suara yang direkam adalah kata “SAYA” selama dua detik seperti pada gambar 3 dibawah ini dengan menggunakan perintah untuk mengambil data rekaman suara dari responden dengan perintah berikut:

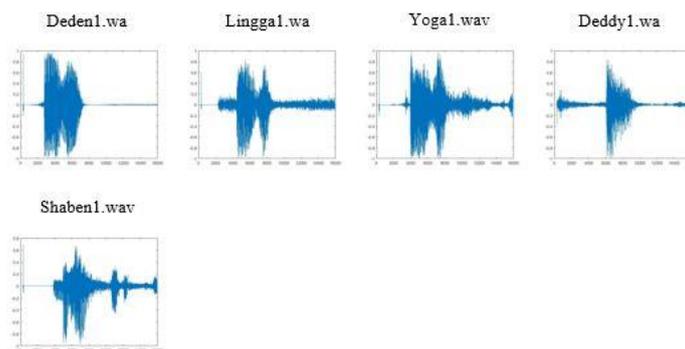
```

y= audioread('sisil.wav');
plot(y);
    
```



Gambar 3. Suara rekaman wanita

Pada gambar 3 merupakan hasil rekaman suara responden wanita, dengan pengucapan kata “saya”, dari hasil rekaman terjadi jeda beberapa detik sebelum pengucapan kata “saya” .



Gambar 4. Suara rekaman pria

Pada gambar 4 merupakan hasil rekaman suara responden pria, dengan pengucapan kata “saya”, dari hasil rekaman terjadi jeda beberapa detik sebelum pengucapan kata “saya”.

### 3.2 Ekstraksi Ciri

Setelah melakukan perekamaan suara, melakukan ekstraksi ciri untuk mendapatkan data yang akurat dan konsisten pada setiap sample, dengan menggunakan metode ekstraksi ciri sinyal suara dengan time domain. Menggunakan rumus (1) untuk ZCR, rumus (2) STE, dan rumus (3) *entropi*. dapat di peroleh hasil seperti pada tabel 1.

Untuk menghitung nilai ZCR rekaman yang digunakan adalah sisi1.wav, rekaman tersebut dimasukan ke variable plot (y). Plot(y) dipanggil untuk menghitung ZCR dengan rumus 1, seperti berikut:

```
s=sign(y);
t=filter([1.1],1,;
zc=(find(t==0))+length(find(y==0));
zcr=length(zc);
```

Hasil dari perhitungan suara sisi1.wav adalah 7209, hasil tersebut di ulang hingga seluruh responden memiliki nilai ZCR.

Hasil perhitungan nilai ZCR di simpan dengan melakukan normalisasi, menghitung STE dengan menggunakan hasil rekaman sisi1.wav untuk mengetahui nilai STE dapat dilihat sebagai berikut:

```
[data, fs] = audioread('tatic1.wav');
% normalize data
data = data / abs(max(data));
```

Hasil dari normalisasi kemudian digunakan untuk menghitung energi dengan rumus 2, dapat dilihat sebagai berikut:

```
%menghitung
energi
energi=(1/length(
data))*sum(abs(da
ta.^2))
```

Hasil dari perhitungan suara sisi1.wav adalah 0.016, hasil tersebut diulang hingga seluruh responden memiliki nilai STE

Hasil perhitungan nilai STE di simpan melanjutkan untuk menentukan tekstur dari suara, menghitung *entropy* dengan rumus 3, menggunakan hasil rekaman sisi1.wav untuk mengetahui nilai *entropy* dapat dilihat sebagai berikut:

```
%menghitung energi frame
%menghitung entropy
Eol = sum (data.^2);
winLength = length(data);
numOfShortBlocks = 1600
%nilai konstanta subframe yang digunakan untuk entropi
subWinlength = floor(winLength / numOfShortBlocks);
```

```
%mencari sub-data;
subdata = reshape(data, subWinlength,numOfShortBlocks)
% menghitung energi sub-fram
s = sum(subdata.^2)/(Eol+eps);
%mencari nilai entropi
entropy = -sum(s.*log2(s+eps));
disp(zcr), disp(energi), disp(entropy)
```

Tabel 1. Hasil Perhitungan Ekstraksi Ciri

NO	Nama	Nama file	Ekstraksi Ciri		
			ZCR	STE	Entropi
1	Sisi	sisi1.wav	7209	0.016	8.128
		sisi2.wav	8384	0.036	8.516
		sisi3.wav	10135	0.016	7.761
2	Tatik	tatik1.wav	6326	0.031	8.761
		tatik2.wav	6455	0.004	8.959
		tatik3.wav	7122	0.041	8.548
...	...	...	...	...	...
6	Deden	deden1.wav	10165	0.0364	8.263
		deden2.wav	7020	0.038	8.418
		deden3.wav	7613	0.031	8.286
7	Lingga	lingga1.wav	3821	0.025	8.362
		lingga2.wav	4488	0.003	8.286
		lingga3.wav	4491	0.012	7.760
...	...	...	...	...	...

Pada tabel 1 merupakan hasil dari perhitungan ZCR, STE dan Entropy untuk menentukan ciri pada masing-masing suara, pada setiap responden menggunakan masing-masing 3 sampel suara. Pada perhitungan ZSC, STE, dan entropy, didapat nilai dari 3 suara, pada suara sisi pada tabel 1 dari 3 sampel suara yg diambil, bahwa suara dari sample 1 hingga sample 3 mengalami peningkatan nilai suara, dan pada STE sampel 2 lebih meningkat. Pada sampel 3 dibagian *entropy* menjadi yang paling rendah. Dari masing-masing ekstraksi ciri pada setiap sampel suara di hitung rata-ratanya seperti pada table 2.

Tabel 2. Perhitungan nilai rata-rata ekstraksi ciri

No	Nama	Nilai rata-rata ekstraksi ciri		
		ZCR	STE	Entropy
1	Sisi	8576	0.022	8.135
2	Tatik	6634	0.024	8.756
3	Nisa	6384	0.020	8.543
4	Yoan	8615	0.031	8.581
5	Wanda	6623	0.064	9.096
6	Deden	8266	0.035	8.322
7	Lingga	4266	0.013	8.136

8	Yoga	5989	0.021	8.297
9	Deddy	5065	0.027	8.302
10	Shaben	6475	0.029	8.405

Pada tabel 2 merupakan hasil dari perhitungan rata-rata ZCR, STE dan Entropy. Hasil tabel 2 terdapat ciri-ciri suara pada masing-masing responden

#### 4. KESIMPULAN

Setelah menghitung rata-rata pada masing-masing suara, dapat di tentukan karakteristik suara yang konsisten dan memiliki ciri suara seperti berikut Rerata wanita ZCR :7366.4, STE: 0.0322, Entropy: 8.6222. Rerata pria ZCR: 6012.2, STE: 0.025, Entropy: 8.2924. Pada hasil rata-rata, nilai ciri suara ZCR tertinggi adalah 8615, dan ciri suara ZCR Terendah adalah 4266. nilai ciri suara STE tertinggi adalah 0.064, dan ciri suara STE Terendah adalah 0.013. nilai ciri suara Entropy tertinggi adalah 9.906, dan ciri suara Entropy Terendah adalah 135.

#### UCAPAN TERIMA KASIH

Dalam kesempatan ini penulis sampaikan rasa terima kasih yang sebesar-besarnya kepada Pusat Penelitian dan Pengabdian Masyarakat Politeknik Negeri Samarinda yang telah memberi dukungan terhadap penelitian ini.

#### DAFTAR PUSTAKA

- [1] "Obtaining permission to use Elsevier material.," p. 99388, 2014, doi: 10.1016/B978-0-08-099388-1.00011-X.
- [2] S. Bagas Bhaskoro and A. Riedho, "Aplikasi Pengenalan Gender Menggunakan Suara," *Semin. Nas. Apl. Teknol. Inf.*, vol. 2012, no. Snati, pp. 15–16, 2012.
- [3] A. Setiawan and P. K. Handayani, "Klastering Suara Berdasarkan Gender dengan Ekstraksi Ciri Berbasis Domain Waktu," *Semin. Nas. Teknol. Inf. dan Komun. Terap. 2012*, vol. 2012, no. Semantik, pp. 364–370, 2012.
- [4] H. Heriyanto and O. S. Simanjuntak, "Identifikasi Ucapan Warna Menggunakan LPC (Linier Predictive Code ) Dan Kelompok Pemilihan Bobot," *Telematika*, vol. 14, no. 01, pp. 68–73, 2017, doi: 10.31315/telematika.v14i01.1968.
- [5] P. Pakan and R. Y. Dillak, "Klasifikasi Musik Menggunakan Polynomial Neural Network," *J. Ilm. Flash*, vol. 3, no. 2, p. 94, 2017, doi: 10.32511/jiflash.v3i2.144.
- [6] C. Altn, "Comparison of Different Time and Frequency Domain Feature Extraction Methods on Elbow Gesture ' s EMG".
- [7] E. S. N. Aisyah, A. W. Hayat, P. Widanti, S. Y. Prasetya, and H. Iskandar, "Analisis Kemiripan Pola Citra Digital Menggunakan Metode Euclidean," *Semin. Nas. Teknol. Inf. dan Multimed.*, pp. 6–8, 2015.
- [8] R. Dimasatria, A. Virgono, and R. R. M. R. Rumani M., "Desain Dan Implementasi Automatic Video Captioning Dengan Speech Recognition Menggunakan Hidden Markov Model," *TEKTRIKA - J. Penelit. dan Pengemb. Telekomun. Kendali, Komputer, Elektr. dan Elektron.*, vol. 1, no. 1, pp. 54–60, 2016, doi: 10.25124/tektrika.v1i1.247.
- [9] T. G. A. PIKRAKIS, *Introduction to A MATLAB Approach AUDIO ANALYSIS* :, 1st ed. The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK 225 Wyman Street, Waltham, MA 02451, USA 525 B Street, Suite 1800, San Diego, CA 92101-4495, USA First: Elsevier, 2014.
- [10] G. Sharma, K. Umapathy, and S. Krishnan, "Trends in audio signal feature extraction methods," *Appl. Acoust.*, vol. 158, p. 107020, 2020, doi: 10.1016/j.apacoust.2019.107020.
- [11] S. Kumar and J. Yadav, "Gender Classification for Emotional Speech using GMFCC and Deep LSTM," no. February, 2020, doi: 10.35940/ijitee.A6109.129219.
- [12] M. Jalil, F. A. Butt, and A. Malik, "Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signals," *2013 Int. Conf. Technol. Adv. Electr. Electron. Comput. Eng. TAECE 2013*, vol. 2, no. 2, pp. 208–212, 2013, doi: 10.1109/TAECE.2013.6557272.
- [13] R. C. Guido, "Knowle dge-Base d Systems ZCR-aided neurocomputing : A study with applications," vol. 105, pp. 248–269, 2016, doi: 10.1016/j.knosys.2016.05.011.
- [14] A. Setiawan, "Analisis Klasifikasi Suara Berdasarkan Gender Dengan Format Wav Menggunakan

- Algoritma K-Means,” *Sains dan Teknol.*, vol. 2, no. 2, pp. 1–8, 2012.
- [15] J. Chen, B. Xu, and X. Zhang, “A Vibration Feature Extraction Method Based on Time-Domain Dimensional Parameters and Mahalanobis Distance,” vol. 2021, 2021.
- [16] S. Anraeni, Herman, and A. Arisandi, “Ekstraksi Fitur Citra Iris Mata Menggunakan Pencirian Matriks Ko-Okurensi Aras Keabuan untuk Klasifikasi Kondisi Kesehatan Ginjal Iris Image Feature Exctraction Using Grey Level Co-Occurance Matrices for Kidney Health Classification,” *Pros. Semin. Nas. Komun. dan Inform.*, vol. 2, no. July, pp. 43–48, 2016.
- [17] “A MATLAB Approach AUDIO ANALYSIS :”, doi: 10.1016/B978-0-08-099388-1.00010-8.
- [18] E. Tarr, *Hack Audio: An Introduction to Computer Programming and Digital Signal Processing in MATLAB*. 2018. doi: 10.4324/9781351018463.